SUPPLEMENT

AN OBJECTIVE METHOD FOR DETERMINING THE PROBABILITY OF TORNADO OCCURRENCE BASED ON ANALYZED WIND SHEAR AND LIFTED INDICES

BY THOMAS M. HAMILL, RUSSELL S. SCHNEIDER, HAROLD E. BROOKS, GREGORY S. FORBES, HOWARD B. BLUESTEIN, MICHAEL STEINBERG, DANIEL MELÉNDEZ, AND RANDALL M. DOLE

This document is supplement B to "The May 2003 Extended Tornado Outbreak," by Thomas M. Hamill, Russell S. Schneider, Harold E. Brooks, Gregory S. Forbes, Howard B. Bluestein, Michael Steinberg, Daniel Meléndez, and Randall M. Dole (*Bull. Amer. Meteor. Soc.*, **86,** 531–542) • ©2005 American Meteorological Society • *Corresponding author:* Dr. Thomas M. Hamill, NOAA–CIRES, Climate Diagnostics Center, Boulder, CO 80305-3328 • E-mail: tom.hamill@noaa.gov • DOI:10.1175/BAMS-86-4-HamillB.

As part of a study of the May 2003 extended outbreak, we sought to objectively determine whether the environmental conditions for this period in May 2003 were unusually conducive for tornado development.

It has long been understood that two crucial ingredients for the generation of supercells that commonly spawn tornadoes are instability and vertical wind shear. Accordingly, National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis data (Kalnay et al. 1996) were used here to examine the statistical relationship between the lifted index (LI), the surface to 500-hPa wind shear, and the probability of strong tornadoes (F2–F5). Here, LIs were used as a surrogate for the more conventional diagnostic, convective available potential energy (CAPE), which is not a reanalysis variable. The goal was to develop a relationship between LI, shear, and tornado probability, and then to examine whether elevated probabilities occurred over an unusually large area for an unusually long time in May 2003.

As an example of the expected relationship, Fig. SB1 provides a sample map of surface to 500-hPa wind shear

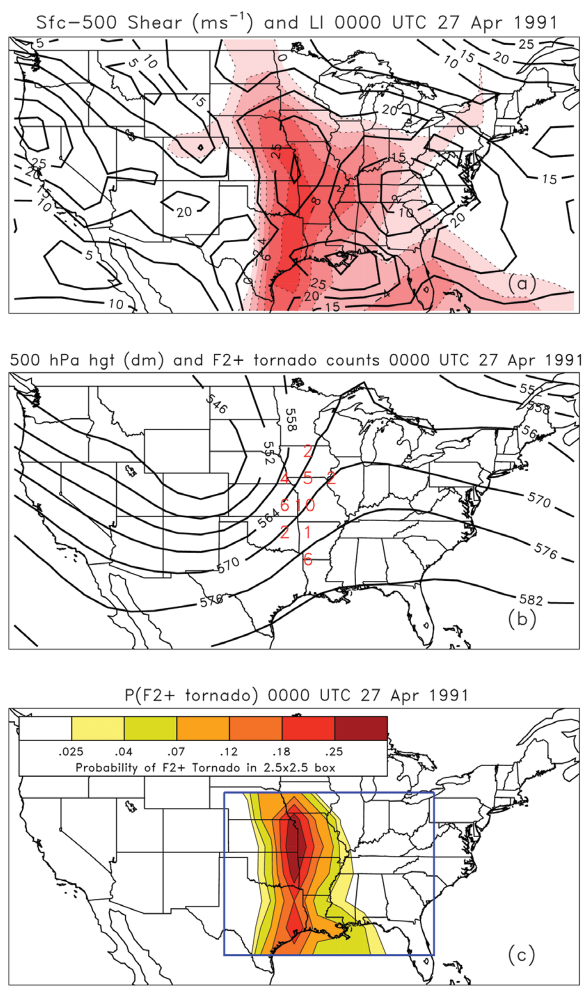

FIG. SB1. (a) Surface to 500-hPa wind shear magnitude and best (lowest four sigma layers) LI from NCEP–NCAR reanalysis on 0000 UTC 27 Apr 1991. (b) The 500-hPa height and number of F2–F5 tornadoes in a 2.5° × 2.5° box for 12 h centered on 0000 UTC. (c) Probability of an F2+ tornado in a 2.5° × 2.5° box for 12 h centered on 0000 UTC, determined from a logistic regression model based on shear and LI (see Fig. SB2 and related discussion).

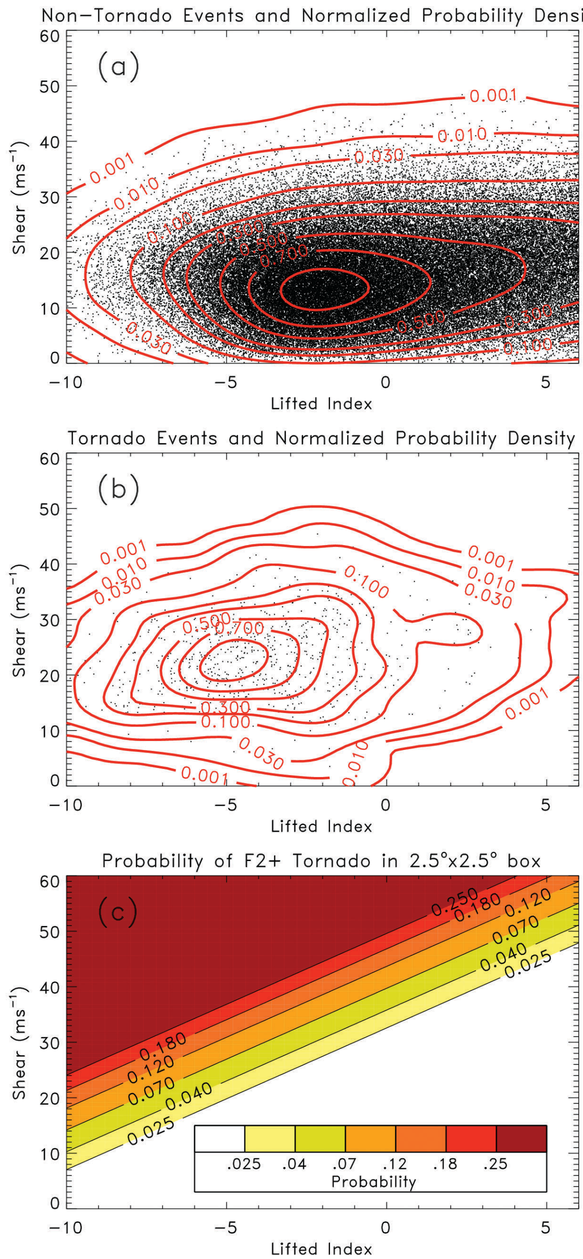

FIG. SB2. (a) Scatterplot of nontornadic on nonsevere tornadic events as a function of LI and shear. Points taken from the box in Fig. 2b of Hammill et al. (2005) for every day in Apr and May from 1979 to 2002. (b) Same as in (a), but for tornadic events. (c) Logistic regression model of tornado probability as a function of shear and LI.

magnitude and "best" LI (lowest computed LI from the bottom four sigma layers), plotted for 0000 UTC 27 April 1991, the day of the Andover, Kansas, tornado outbreak. Also plotted are 500-hPa height and the number of F2–F5 tornadoes in the 12 h centered on 0000 UTC. As expected, the strong tornadoes occurred in or nearby a region where the shear was high and the LI was low.

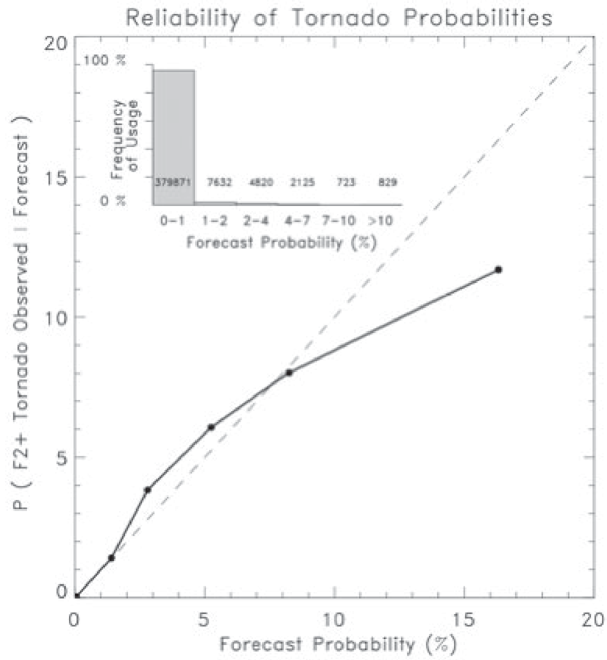

FIG. SB3. Reliability of tornado "forecasts" based on analyzed shear and LI. Inset histogram provides the frequency of issuance for each of the six binned forecasts, indicating that most of the time less than 1% probability is issued (379,871 forecasts with such a probability), but extreme probabilities are issued only very rarely (829 total). Data from Apr–May 1979 to 2002 using grid boxes inside blue box in Fig. SB1c.

Next, the relationship between tornado occurrence and wind shear and instability was quantified. Inside an area over the Midwest (the blue box in Fig. SB1c), each 2.5° × 2.5° grid box was either classified as strong tornadic (at least one F2+ tornado occurred between 1800 and 0600 UTC) or nontornadic (no F2+ occurred) for each day in April and May from 1979 to 2002. The shear and LI were also noted at these grid points. Figure SB2 provide a scatterplot of the nontornadic samples (Fig. SB2a) and the tornadic samples (Fig. SB2b). An estimate of the normalized probability density is also plotted over the data. As expected, tornadoes more commonly occur at higher shears and lower LIs than do nontornadic events. The maximum probability density for nontornadic events was at approximately LI = −2 and shear = 13 m s^{-1}, while the maximum density for tornadic events was at ~ LI = −5 and shear = 22 m s^{-1}. There were some occurrences of tornadoes at positive LIs. These may be cases where the mesoscale environment at the time of the tornado was less stable than the large-scale 0000 UTC analysis. Another possibility is a poor analysis; the humidity and surface temperature that are necessary for the LI calculation are notoriously difficult to analyze correctly.

Nonetheless, the data appeared to support the hypothesis that analyzed shear and LI might be able to be used to discriminate between tornadic and nontornadic conditions. Accordingly, we developed a logistic regression model (Wilks 1995) to predict the conditional probability of a F2+ tornado given the analyzed 0000 UTC shear and LI. The logistic regression model is of the form $P(\text{Tornado}|\text{LI, shear}) = 1.0 - 1.0/\{1.0 + \exp[b_0 + (b_1 \times \text{shear}) + (b_2 \times \text{LI})]\}$. Using the 1979–2002 data (without cross validation), the logistic regression was used to predict the regression coefficients b_0, b_1, and b_2. The resulting probability model of tornado probabilities as a function of shear and LI is shown in Fig. SB2c. As expected, tornado probabilities increase as LI decreases and shear increases. Figure SB1c illustrates the predicted probability of tornadoes on the day of the Andover, Kansas, tornado outbreak. Figure SB3 shows that the tornado probabilities are quite reliable.

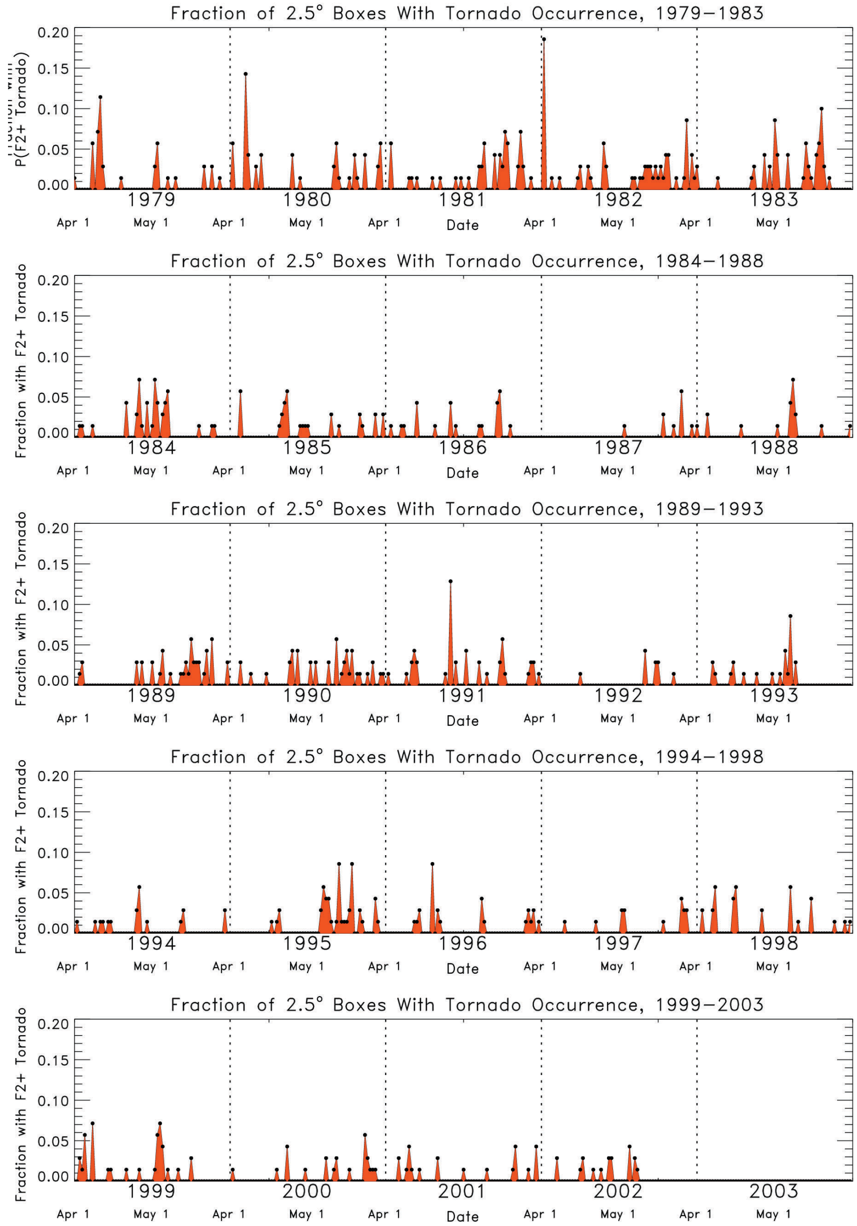

FIG. SB4. As in Fig. 9 in Hamill et al. (2005), but for the fraction of grid boxes that had F2+ tornadoes. (Data for 2003 not available.)

While section 2 of the Hamill et al. (2005) article demonstrates that the May 2003 tornado outbreak was unusually large and sustained, it would be useful to quantify whether the environmental conditions supporting tornado development were anomalously conducive over a large region for a long time during May 2003. This can now be evaluated using the probability model. Let us choose a probability threshold of 2.5% as indicative of a significantly elevated risk of a strong tornado, which is approximately 3.5 times larger than the climatological probability. The fraction of grid points that have this elevated risk (inside the blue box in Fig. SB1c) were calculated, and a time series of this fraction was plotted (Fig. 9 in Hamill et al. 2005). The time series shows that since 1979 no other period other than early May 2003 had a string of nine straight days where the fractional coverage of elevated risk was greater than 0.2. This week in May 2003 did indeed have a sequence of days where shear and LI were climatologically anomalous over a large region. But the time series also suggests that periods of several days of sustained tornado-favorable conditions may be more frequent in recent years than they were back in the 1980s.

However, the notion that sustained tornado outbreaks are happening more frequently is not supported by the observational data. Figure SB4 provides

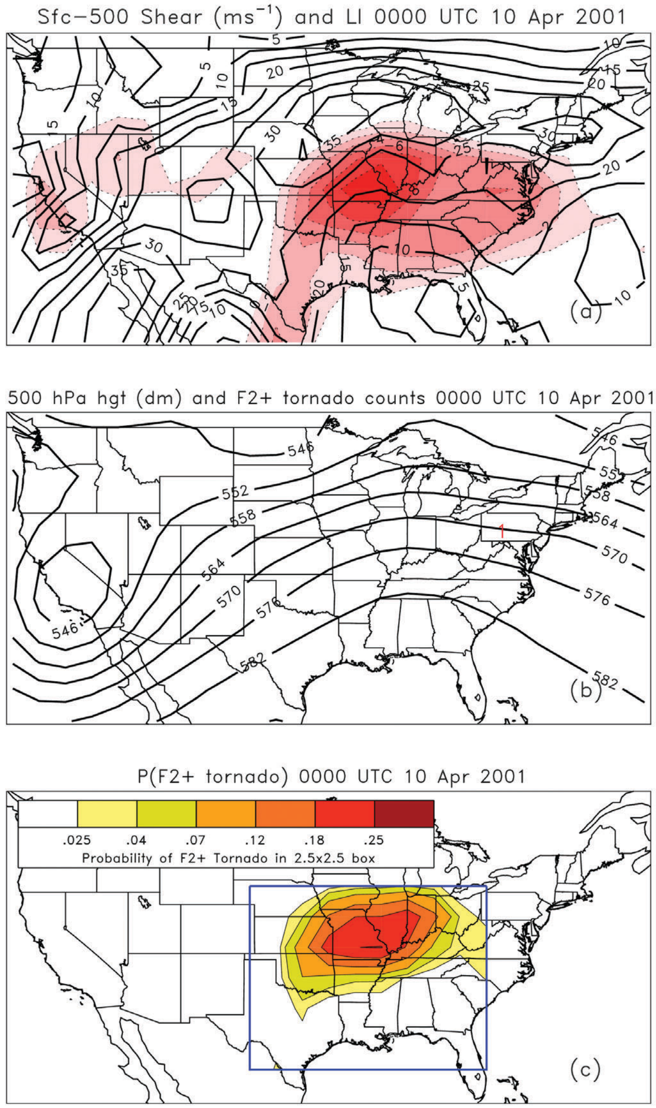

Fig. SB5. As in Fig. SB1, but for 0000 UTC 10 Apr 2001.

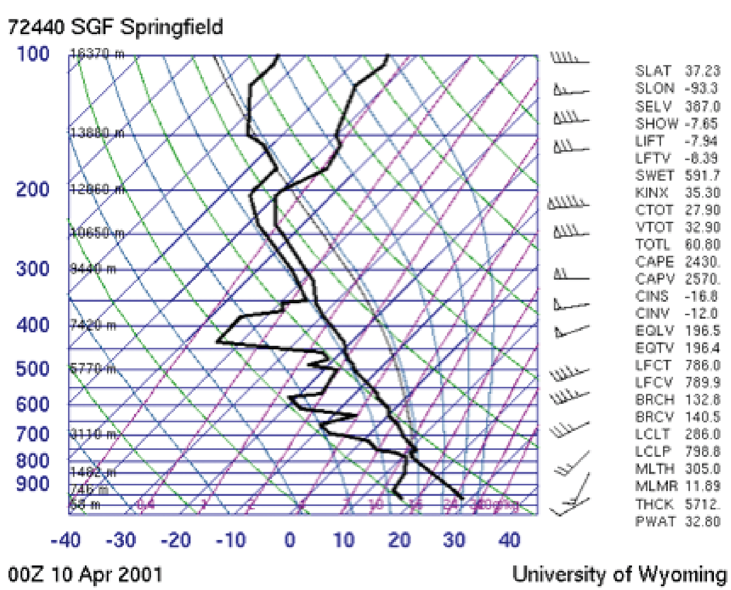

Fig. SB6. Skew *T*–log*p* from Springfield, MO, on 0000 UTC 10 Apr 2001. (Courtesy of the University of Wyoming Atmospheric Sciences Web site.)

a time series of the fraction of 2.5° × 2.5° grid boxes that had strong tornadoes. This time series of actual tornadoes appears to be more stationary, with no evidence that the areal extent or frequency of severe tornado outbreaks is increasing. Note by comparing the two time series that the large outbreaks generally correspond to situations where the fractional coverage of high tornado probability was large, but there are also frequent occurrences where a large area of high probability existed but tornadoes did not occur (the two time series correlate at 0.45). Figure SB5 provides an example of such a "false alarm." Tornado probabilities for this case day were high in southern Missouri, but tornadoes did not occur. It does not appear that the errors are due to a faulty reanalysis; Skew *T*–log*p* data at this time (Fig. SB6) support the analysis of large shear and low LIs. Obviously, there are more factors that play a role in tornadogenesis than merely adequate shear and instability, such as the presence or absence of mesoscale boundaries, convective inhibition, etc.

REFERENCES

Hamill, T. M., R. S. Schneider, H. E. Brooks, G. S. Forbes, H. B. Bluestein, M. Steinberg, D. Meléndez, and R. M. Dole, 2005: The May 2003 extended tornado outbreak. *Bull. Amer. Meteor. Soc.,* **86,** 531–542.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–472.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.